Prof. Dr. Peter von der Lippe

Modelle für Paneldaten

Eine kurze Einführung in die Panelökonometrie

* als spezielles Modell mit objektspezifischem intercept

Die Darstellung ist stark angelehnt an Vorlesungsunterlagen von Prof. Dr. Eckey (Univ. Kassel). Zur Interpretation der Modelle wurden auch herangezogen M. P. Murray, Econometrics, Stock and Watson und Greene, Introduction to Econometrics (beide Pearson, Addison Wesley Series).[1]

1. Gegenstand, Notation, Daten

Panelanalysen stellen eine Kombination von Querschnittsanalyse mit N Einheiten, Individuen (entities) oder – hier bevorzugt – "Objekten" (units, entities), etwa N = 16 Bundesländer[2] oder N = 10 Branchen und Längsschnittsanalyse (besser Zeitreihenanalyse) über T Perioden dar. Bei einem Panel wird vorausgesetzt, dass sich die Daten (für die Variablen y, x_2, …, x_k) im Zeitablauf ($\forall$ t = 1, …, T) *auf die gleichen Einheiten beziehen*, andernfalls spricht man von repeated cross sections[3]. Idealerweise gibt es bei N Objekten und T Perioden keine fehlenden Angaben (der Datensatz ist dann "balanced", andernfalls [bei Unvollständigkeit] "unbalanced").

Notation

Subskripte	Variablen	Parameter
j = 1, …, N Einheiten[a)] (Objekte)	y_{jt} endogene Variable	β_{ji} Regressionskoeffizienten
t = 1, …, T Perioden (Zeitreihe)	x_{ijt} exogene Variable	β_{j1} (= α_j) intercept
i = 2, …, k Regressoren (k = K+1)	u_{jt} Störterm	

a) Wenn j von einem anderen Objekt unterschieden werden soll wird auch der Index m benutzt (j ≠ m).

Struktur der Daten (der "Pool" in spread sheet view)

	t	j = 1	j = 2	…	j = N	
		Pool-Datensatz				**Panel**-Datensatz
y_{jt}	t = 1	y_{11}	y_{21}	…	y_{N1}	Spaltenvektoren untereinander
	…	…	…	…		$\mathbf{y}_1\ \mathbf{x}_{21}\ \mathbf{x}_{31}$
	t = T	y_{1T}	y_{2T}	…	y_{NT}	$\mathbf{y}_2\ \mathbf{x}_{22}\ \mathbf{x}_{32}$
x_{2jt}	t =1	x_{211}	x_{221}	…	x_{2N1}	…
	…	…	…	…	…	$\mathbf{y}_N\ \mathbf{x}_{2N}\ \mathbf{x}_{3N}$
	t = T	x_{21T}	x_{22T}	…	x_{2NT}	
x_{3jt}	t = 1	x_{311}	x_{321}	…	x_{3NT}	

[1] Für eine kritische Durchsicht des Textes sowie wichtige Anregungen und Verbesserungsvorschläge danke ich Herrn J. Mehrhoff (Deutsche Bundesbank).

[2] Bei Eckey wurde versucht, die unterschiedliche Größe der Bundesländer durch geeignete Zusammenfassung mehrerer Bundesländer zu etwa gleich großen Bundesländergruppen auszugleichen. Hierzu besteht keine Notwendigkeit. Es kann im Gegenteil sogar gewünscht sein, die Unterschiedlichkeit der Bundesländer hinsichtlich ihrer Größe (Einwohnerzahl) in die Betrachtung einzubeziehen, weil sie – neben anderen Faktoren – eine Ursache für die sich in unterschiedlichen α Koeffizienten ausdrückende "unobserved heterogeneity" sein kann.

[3] Der Sprachgebrauch ist anders als in der Wirtschaftsstatistik und Umfrageforschung. Panel im obigen Sinne ist danach eine echte Längsschnittsanalyse bzw. Kohortenanalyse (longitudinal analysis).

Der gelb markierte Bereich ist der T ×1 Spaltenvektor $\mathbf{y}_1$ der türkis markierte ist der Vektor ist $\mathbf{x}_{2N}$ usw. Bei EViews stehen diese Daten jeweils nicht untereinander sondern nebeneinander. Man findet die untereinander geschriebenen (zu NT ×1 Blockvektoren) zusammengefassten Vektoren $\mathbf{y}_1, \ldots, \mathbf{y}_N$ und entsprechend $\mathbf{x}_{21}, \mathbf{x}_{22}, \ldots, \mathbf{x}_{2N}, \mathbf{x}_{31}, \ldots$ usw. auch unten in Gl. 3.

2. Getrennte Schätzung für jedes Objekt (nicht-restringiertes Modell)

Die Regressionsgleichung für das Objekt j zur Zeit t lautet bei i = 1, …, k Regressoren

(1) $$y_{jt} = \beta_{j1} + \beta_{j2} x_{2jt} + \ldots + \beta_{jk} x_{kjt} + u_{jt}$$

Man beachte die Änderung der Notation: das Absolutglied (intercept) ist jetzt β_{j1} statt α_j, $x_{1t} = 1$ $\forall$ t und es gibt K = k–1 "echte" Regressoren $x_{2t}, x_{3t}, \ldots, x_{kt}$.

Für das j-te Objekt hat man somit folgende Daten (der gesamte Längsschnitt mit T Werten für y, $x_2, x_3, \ldots, x_k$)

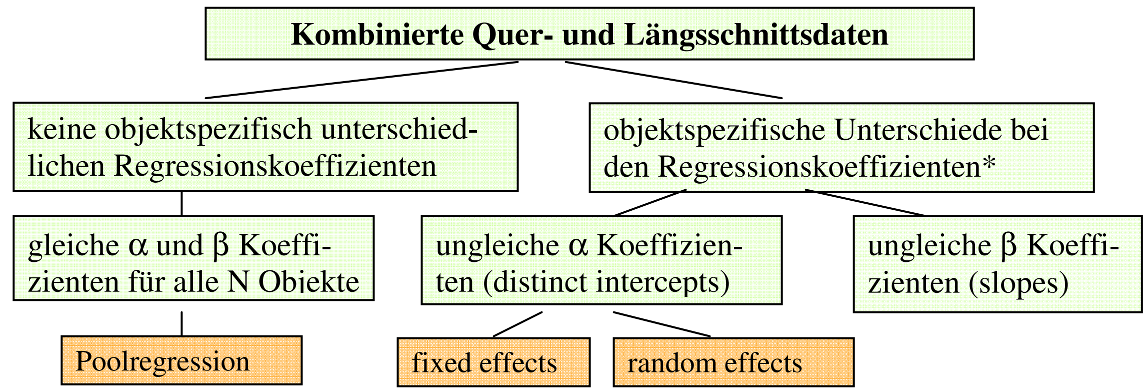
$$\mathbf{y}_j = \begin{bmatrix} y_{j1} \\ y_{j2} \\ \vdots \\ y_{jT} \end{bmatrix}, \mathbf{X}_j = \begin{bmatrix} 1 & x_{2j1} & \cdots & x_{kj1} \\ 1 & x_{2j2} & \cdots & x_{kj2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{2jT} & \cdots & x_{kjT} \end{bmatrix} \text{ und } \boldsymbol{\beta}_j = \begin{bmatrix} \beta_{j1} \\ \beta_{j2} \\ \vdots \\ \beta_{jk} \end{bmatrix} = \begin{bmatrix} \alpha_j \\ \beta_{j2} \\ \vdots \\ \beta_{jk} \end{bmatrix} \text{ sowie } \mathbf{u}_j = \begin{bmatrix} u_{j1} \\ u_{j2} \\ \vdots \\ u_{jT} \end{bmatrix}$$

(1a) $\mathbf{y}_j = \mathbf{X}_j \boldsymbol{\beta}_j + \mathbf{u}_j$ für j = 1, …, N (Gl. 1 ist die t-te Zeile dieses Gleichungssystems)

Die OLS Schätzung des j- ten Objekts ist

(2) $$\hat{\boldsymbol{\beta}}_j = \left(\mathbf{X}_j'\mathbf{X}_j\right)^{-1}\mathbf{X}_j'\mathbf{y}_j$$

und man erhält als Summe der Quadrate der Abweichungen $\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j = \sum_{t=1}^{T} \hat{u}_{jt}^2$. Schätzt man alle N Regressionsgleichungen nach Art von Gl. 2 so erhält man N Koeffizientenvektoren $\hat{\boldsymbol{\beta}}_1, \ldots, \hat{\boldsymbol{\beta}}_N$ vom Typ k × 1 sowie N Vektoren $\hat{\mathbf{u}}_j$ der geschätzten Störgrößen (jeweils Vektoren vom Typ T × 1).[4] Dieses Modell "Getrennte Regressionen" ist nicht-restringiert (constrained) und im Vergleich dazu sind andere in der folgenden Übersicht genannte Modelle spezieller:



* weitere mögliche Unterschiede: time effects (modelliert mit time dummies)

[4] Man beachte, dass an späterer Stelle auch Vektoren $\boldsymbol{\beta}$ eingeführt werden, bei denen der α Koeffizient nicht vorkommt, die also insofern "reduziert" sind. Um sie von $\boldsymbol{\beta}$ zu unterscheiden werden sie mit $\boldsymbol{\beta}^*$ bezeichnet.

Eine wichtige Größe, die im Folgenden in verschiedenen Teststatistiken vorkommt (vgl. Gl. 5, 12 und 14) ist $\sum_{j=1}^{N}\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j = \sum_{j=1}^{N}\sum_{t=1}^{T}\hat{u}_{jt}^2$, die Summe der Abweichungsquadrate (sum of squared residuals) im nicht-restringierten Modell. In anderen Modellen werden (gemäß H_0) Einschränkungen gemacht für die Koeffizienten (d.h. sie sind "restringiert"):

Modell	Restriktion (H_0)
Getrennte Regressionen[1]	Keine Restriktionen für die Regressionskoeffizienten
Poolregression[2]	β_{j1} (= α_j), β_{j2}, …, β_{jk} für alle N Objekte gleich, also $\boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_N = \boldsymbol{\beta}$
feste und zufällige Effekte	unterschiedliche α Koeffizienten (α_1, …, α_N) aber gleiche β Koeffiz.
unterschiedliche β Koeff.	gleiche α Koeffizienten ($\alpha_1 = \ldots = \alpha_N = \alpha$) aber unterschiedl. β Koeffiz.

1) gar nicht restringiert
2) am stärksten restringiert

3. Poolregression (gepooltes Modell) (gleiche Regressionskoeffizienten α und β für alle N Objekte)

Die Koeffizienten $\beta_1 = \alpha$, β_2, …, β_k sind für alle T Perioden und alle N Einheiten gleich (konstant), so dass gilt $\boldsymbol{\beta}_j = \boldsymbol{\beta} = \begin{bmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_k\end{bmatrix}$ $\forall j$. Die Datenvektoren werden im Prinzip aufeinandergestapelt (stacked) etwa[5] $\mathbf{y} = \begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \vdots\\ \mathbf{y}_N\end{bmatrix}$ mit $\mathbf{y}' = [y_{11} \; \ldots \; y_{1T} \; y_{21} \; \ldots \; y_{2T} \; \ldots \; y_{N1} \; \ldots \; y_{NT}]$ ferner hat man die genauso dimensionierten Vektoren[6] $\mathbf{u}$ der wahren und $\hat{\mathbf{u}}_p$ der geschätzten Störgrößen.[7] Das Modell $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ ist jetzt im Fall der Poolregression voll ausgeschrieben

$$\text{(3)} \qquad \underset{NT\times 1}{\mathbf{y}} = \begin{bmatrix} y_{11}\\ \vdots\\ y_{1T}\\ \vdots\\ y_{N1}\\ \vdots\\ y_{NT}\end{bmatrix} = \begin{bmatrix} 1 & x_{211} & \cdots & x_{k11}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_{21T} & \cdots & x_{k1T}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & x_{2N1} & \cdots & x_{kN1}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & x_{2NT} & \cdots & x_{kNT}\end{bmatrix} \cdot \begin{bmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_k\end{bmatrix} + \begin{bmatrix} u_{11}\\ \vdots\\ u_{1T}\\ \vdots\\ u_{N1}\\ \vdots\\ u_{NT}\end{bmatrix} = \underset{NT\times k}{\mathbf{X}} \cdot \underset{k\times 1}{\boldsymbol{\beta}} + \underset{NT\times 1}{\mathbf{u}}$$

und die OLS Schätzung (z. B. mit EViews) des ***gepoolten*** Modells[8]

---

[5] Man kann sagen **y** ist ein NT × 1 Blockvektor, d. h. jede Zeile $\mathbf{y}_1$, $\mathbf{y}_2$, …, $\mathbf{y}_N$ ist selbst wieder ein Vektor, oder anders gesagt y ist partitioned by rows.
[6] Auch hier wieder Blockvektoren: N übereinander gestapelte Spaltenvektoren mit T Zeilen.
[7] Weitere Betrachtungen unter Verwendung von Blockmatrizen und Blockvektoren finden sich im Anhang.
[8] Die folgende Gleichung bezieht sich auf alle N Objekte, nicht nur auf das j-te wie bei Gl. 2.

(4) $\hat{\boldsymbol\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$

liefert den $k \times 1$ Vektor $\hat{\boldsymbol\beta}$ sowie $S^0_{\hat u\hat u} = \hat{\mathbf{u}}_p'\hat{\mathbf{u}}_p$ als "Sum squared resid". Es ist wichtig, diese Größe zu unterscheiden von der Summe der Residuenquadrate bei N ***getrennten*** Schätzungen $S_{\hat u\hat u} = \sum_{j=1}^{N}\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j$ (siehe oben Abschn. 2).

Die F-verteilte Prüfgröße der Hypothese der Gleichheit aller Regressionskoeffizienten (kein Unterschied zwischen den N Objekten) also zur Prüfung von H_0: $\boldsymbol\beta_1 = \boldsymbol\beta_2 = \ldots = \boldsymbol\beta_N = \boldsymbol\beta$ ist dann

(5) $$F_p = \frac{\left(\hat{\mathbf{u}}_p'\hat{\mathbf{u}}_p - \sum_{j=1}^{N}\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j\right) \Big/ [(N-1)k]}{\sum_{j=1}^{N}\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j \Big/ (NT - Nk)}.$$

Es ist intuitiv verständlich, dass $F_p = 0$ gilt, wenn H_0 zutrifft, denn dann sind (da die Restriktionen zutreffen) $\sum_{j=1}^{N}\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j$ und $\hat{\mathbf{u}}_p'\hat{\mathbf{u}}_p = \begin{bmatrix}\hat{\mathbf{u}}_{p1}' & \cdots & \hat{\mathbf{u}}_{pN}'\end{bmatrix}\begin{bmatrix}\hat{\mathbf{u}}_{p1} \\ \vdots \\ \hat{\mathbf{u}}_{pN}\end{bmatrix}$ identisch, so dass der Zähler von F_p Null ist. Die Quadratsumme der geschätzten Residuen bei Geltung von H_0 also $\hat{\mathbf{u}}_p'\hat{\mathbf{u}}_p$ entspricht[9] $S^0_{\hat u\hat u}$ und entsprechend ist $\sum_{j=1}^{N}\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j$ (im Modell der getrennten Regressionen) das, was bisher $S_{\hat u\hat u}$ genannt wurde.

4. Modelle mit objektspezifischen α Koeffizienten (distinct intercepts $\alpha_1, \ldots, \alpha_N$)

Die Individualität eines Objekts j im Verhältnis zu einem anderen Objekt m drückt sich in unterschiedlichen α Koeffizienten aus mit $\alpha_j \neq \alpha_m$ (j, m = 1, …, N, j ≠ m) im Unterschied zu einem für alle N Objekte gleichen Koeffizienten α im Modell der Pool-Regression. Man kann sich α_j wie folgt zusammengesetzt denken

(6) $\alpha_j = \alpha + \mu_j$

mit $E(\mu_j) = 0$ als Ausdruck einer "unobserved heterogeneity" (alle über die Zeit konstante aber für das jeweilige Objekt spezifischen, nicht beobachtbaren Einflüsse).[10] Für das Objekt j zur Zeit t lautet das Modell dann (mit $\alpha = \beta_1$)

(7) $y_{jt} = \alpha_j + \beta_2 x_{2jt} + \ldots + \beta_k x_{kjt} + u_{jt} = (\alpha + \mu_j) + \beta_2 x_{2jt} + \ldots + \beta_k x_{kjt} + u_{jt}$.

Hinsichtlich der Annahmen über μ_j wird unterschieden:

- **fixed effects** in der Grundgesamtheit sind α und $\mu_1, \ldots, \mu_N$ konstante Größen[11] (die Schätzwerte $\hat\alpha_1,\ldots,\hat\alpha_N$ sind natürlich, wie immer, so auch hier, Zufallsvariabeln)

und

[9] Weil das gepoolte Modell ja insofern "restringiert" ist, als angenommen wurde, dass die Regressionskoeffizienten nicht nur konstant sind über die Zeit, sondern auch für alle N Objekte gleich sind.

[10] Wird eine bestehende unobserved heterogeneity vernachlässigt (also fälschlich H_0: $\alpha_1 = \ldots = \alpha_N = \alpha = \beta_1$ angenommen obgleich dies nicht gilt), so entsteht ein omitted variables bias.

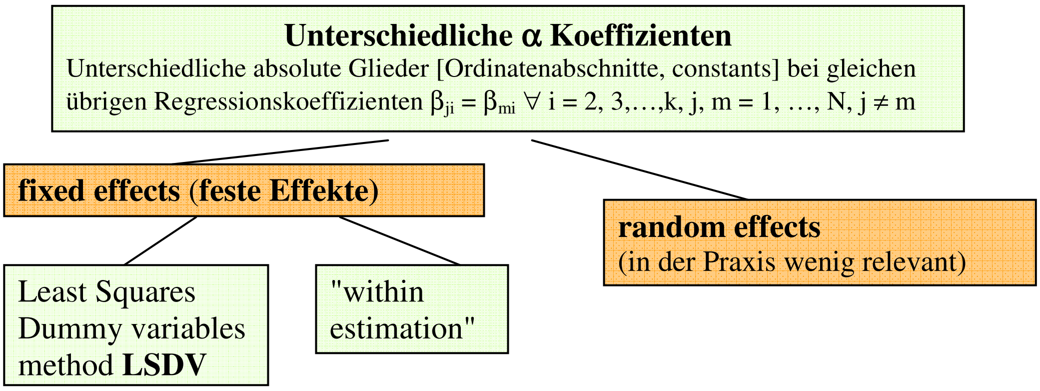
[11] Und damit natürlich auch $\alpha_1 = \alpha + \mu_1, \ldots, \alpha_N = \alpha + \mu_{1N}$.

- **random effects**, wobei die Größen $\mu_1, \dots, \mu_N$ (auch in der Grundgesamtheit) zufällig schwanken (Zufallsvariablen sind), so dass man die Störgröße u_{jt} mit μ_j zu einer Zufallsvariable $v_{jt} = \mu_j + u_{jt}$ zusammenfassen kann (über die Verteilung von v_{jt} und über die Kovarianzen von v_{jt} mit x_{ijt} werden Annahmen gemacht), so dass statt Gl. 7 gilt

(7a) $$y_{jt} = \alpha + \beta_2 x_{2jt} + \dots + \beta_k x_{kjt} + (\mu_j + u_{jt}) = \alpha + \Sigma\beta_i x_{ijt} + v_{jt}$$

mit $v_{jt} = \mu_j + u_{jt}$.

Für das Modell mit fixed effects sind zur Schätzung der Koeffizienten β_i (i = 2, …, k), also des reduzierten (nicht mehr $\alpha = \beta_1$ enthaltenden) $K \times 1$ ($K = k-1$) Parametervektors $\boldsymbol{\beta}^* = \begin{bmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix}$ zwei Verfahren üblich, die in der folgenden Übersicht genannt sind und auf die in Abschn. 5 eingegangen wird.



Beim random effects Modell hat die Störgröße v_{jt} (total disturbance) zwei Komponenten (sie ist eine Summe von zwei Zufallsvariablen):[12]

- die "individual error component" μ_j, für die üblicherweise $E(\mu_j) = 0$ angenommen wird, so das $E(\alpha_j) = \alpha$ = const. gilt, d. h. die α_j zufällig um α schwanken;[13]
- die zeit- und objektabhängigen Zufallsvariablen u_{jt} (j = 1, …, N und t = 1, …, T).

Es gibt erheblich mehr Schätzprobleme beim random effects model als beim fixed effects model, weil die Autokovarianz[14] von v_{jt} und v_{jt*} (wegen der Gemeinsamkeit von μ_j über alle t) auch unter den im Folgenden aufgelisteten Annahmen nicht verschwindet.[15]

Hinsichtlich μ_j und u_{jt} (und damit auch über $v_{jt} = \mu_j + u_{jt}$) sind folgende Annahmen üblich[16]

[12] Man (z. B. Murray) spricht deshalb auch vom "**error components**" Modell (synonym mit random effects).
[13] μ_j ist eine *zufällige* objektspezifische Abweichung des Objekts j vom allgemeinen Durchschnitt. Für das Objekt m gilt entsprechend $v_{mt} = \mu_m + u_{mt}$.
[14] Mit t, t* = 1, …, T und t ≠ t*.
[15] Vielmehr gilt nach Murray, S. 691: $E(v_{jt}v_{jt*}) = E(\mu_j + u_{jt})(\mu_j + u_{jt*}) = V(\mu_j) = V(\mu)$. Es sind Schätzungen der Varianz von μ nötig, und man spricht hier (wie auch sonst bei Heteroskedastizität und/oder Autokorrelation) von der geschätzten verallgemeinerten Methode der kleinsten Quadrate (= feasible generalized least squares FGLS).
[16] Nach Murray, S. 683.

	individual error component μ_j	zeit- und objektabhängigen Zufallsvariable u_{jt}
Standardannahmen	$E(\mu_j) = 0,\ V(\mu_j) = \sigma^2_\mu$	$E(u_{jt}) = 0,\ V(u_{jt}) = \sigma^2_u\ (\forall\ j, t)$
uncorrelated acrosss individuals	$E(\mu_j\mu_m) = 0$	$E(u_{jt}u_{mt*}) = 0$ für $j \neq m$ und $t \neq t^*$
Die μ_j sind nicht mit den u_{jt} korreliert	$E(\mu_j u_{mt}) = 0\ (\forall\ j, m, t)$	

Neben den Standardannahmen[17] bezüglich der Störgrößer u_{jt} ist auch die Annahme der Exogenität aller Regressoren x_{ijt} wichtig.[18] $E(x_{ijt}v_{jt}) = 0\ (\forall\ i, j, t)$ kann nicht ohne weiteres angenommen werden, denn μ_m (bzw. μ_j) ist nicht notwendig für jede Periode t "uncorrelated with explanators" für x_{ijt} ($i = 2, \ldots, k$).

Gilt $E(x_{ijt}\mu_m) = 0$, ist also die objektspezifische Zufallsvariable μ_j (die individuelle Fehlerkomponente) mit den Regressoren (explanators) x_{ijt} unkorreliert, dann liegt das (speziellere) random effects Modell vor,[19] andernfalls (wenn $E(x_{ijt}\mu_m) \neq 0$) das Modell der fixed effects.

Im random effects Modell bietet sich eine Schätzung mit **feasible generalized least squares (FGLS)** an. Hierzu ist zusätzlich zu σ^2_u in der Matrix $\mathbf{\Omega} = E(\mathbf{vv}')$ der Parameter $V(\mu) = \sigma^2_\mu$ (gerechnet über $j = 1, \ldots, N$) zu schätzen, denn bei FGLS ist die folgende $T \times T$ Matrix (Kovarianzmatrix der Störgrößen) $\mathbf{\Omega}$ zu bestimmen[20]

$$\text{(8)} \qquad \mathbf{\Omega} = \begin{bmatrix} \sigma^2_u + \sigma^2_\mu & \sigma^2_\mu & \cdots & \sigma^2_\mu \\ \sigma^2_\mu & \sigma^2_u + \sigma^2_\mu & \cdots & \sigma^2_\mu \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2_\mu & \sigma^2_\mu & \cdots & \sigma^2_u + \sigma^2_\mu \end{bmatrix} = \sigma^2_u \mathbf{I}_T + \sigma^2_\mu \mathbf{i}_T \mathbf{i}'_T \,.$$

Man kann $\mathbf{\Omega}^{-1/2}$ auch als Transformationsmatrix interpretieren, mit der die y und x Werte so aufeinander projiziert werden, dass die Kovarianzmatrix der transformierten Störgröße wieder die Effizienzbedingungen der OLS-Methode erfüllt (Homoskedastizität und Freiheit von Autokorrelation).

das "wahre" Modell der Grundgesamtheit	Schätzverfahren		
	OLS	LSDV*	FGLS
pool a)	effizient	konsistent b)	konsistent
fixed effects	inkonsistent c)	effizient	inkonsistent d)
random effects	konsistent	konsistent	effizient e)

* Least squares dummy variables (man beachte, dass LSDV nie inkonsistent sein kann)
a) Alle α Koeffizienten in der Grundgesamtheit gleich, daher $V(\mu) = 0$.
b) Weniger effizient als FGLS (es gehen N–1 Freiheitsgrade wegen der N–1 Dummies verloren) während bei FGLS wegen der Schätzung von $V(\mu)$ nur ein Freiheitsgrad verloren geht.
c) Wegen omitted variables bias.
d) Weil die nicht dem wahren Modell entsprechende Störgröße nicht uncorrelated with explanators (x_{ijt}) ist.
e) Gilt für GLS nicht FGLS (nur konsistent).

[17] B1 bis B3 im Lehrbuch von L. v. Auer (also $E(u_{.t}) = 0$, Homoskedastizität und keine Autokorrelation).
[18] Das ist Annahme C1 ($E(x_{ijt}u_{jt}) = 0$ für alle k Regressoren $i = 1, \ldots, k$) bei von Auer.
[19] Zum Testen der Annahme der Unkorreliertheit – und damit der Konsistenz einer OLS Schätzung – ist der **Hausman Spezifikationstest** [wie auch bei der Annahme C1] anzuwenden; vgl. Murray, S.701.
[20] Greene, S. 568.

Auf der Hauptdiagonale (farblich markiert) findet man die Kombinationen, in denen Modell und Schätzverfahren jeweils genau aufeinander abgestimmt sind (d. h. es wird das für den betrachteten Fall entwickelte Schätzverfahren angewendet).

5. Schätzung der Parameter in Modellen mit distinct intercepts α

Sind nur die α Koeffizienten ungleich (objektspezifisch), so lässt sich der Parametervektor $\boldsymbol{\beta}^{*}{}' = [\beta_2 \quad \beta_3 \quad \cdots \quad \beta_k]$ mit LSDV oder "within estimation" schätzen.[21] Wie obige Übersicht zeigt, hat LSDV Vorteile, weil das Verfahren selbst dort, wo es vom Modell her nicht angebracht ist, nie inkonsistent ist.

5.1. Schätzung mit dem "within estimation" Verfahren

Es empfiehlt sich, von allen Variablen jeweils den (über die Zeit, also alle T Perioden gerechneten) Mittelwert abzuziehen. Geht man aus von

(1a) $\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{u}_j$ bzw. bei nicht objektspezifischen β Koeffizienten

(1b) $$\mathbf{y}_j = \begin{bmatrix} y_{j1} \\ \vdots \\ y_{jT} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \cdot \alpha_j + \begin{bmatrix} x_{2j1} & \cdots & x_{kj1} \\ x_{2j2} & \cdots & x_{kj2} \\ \vdots & \ddots & \vdots \\ x_{2jT} & \cdots & x_{kjT} \end{bmatrix} \cdot \begin{bmatrix} \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_{j1} \\ \vdots \\ u_{jT} \end{bmatrix} \text{ oder } \mathbf{y}_j = \mathbf{i}\alpha_j + \mathbf{X}_j^{*}\boldsymbol{\beta}^{*} + \mathbf{u}_j$$

worin $\boldsymbol{\beta}^{*}$ der für alle N Objekte (j = 1, …, N) gleiche (es gilt also $\boldsymbol{\beta}_1^{*} = \ldots = \boldsymbol{\beta}_N^{*} = \boldsymbol{\beta}^{*}$) um $\beta_{1j} = \alpha_j$ reduzierte[22] Parametervektor, um dessen Schätzung es hier geht, und $\mathbf{i}\alpha_j = \begin{bmatrix} \alpha_j \\ \vdots \\ \alpha_j \end{bmatrix}$ ein $T \times 1$ Vektor ist. Werden die Variablen jeweils vom Mittelwert bereinigt (zentriert oder "demeaned", was hier durch die Tilde ~ gekennzeichnet werden soll), dann gilt mit

$$\tilde{\mathbf{y}}_j = \begin{bmatrix} y_{j1} - \bar{y}_j \\ \vdots \\ y_{jT} - \bar{y}_j \end{bmatrix} \text{ und } \tilde{\mathbf{X}}_j^{*} = \begin{bmatrix} x_{2j1} - \bar{x}_{2j} & \cdots & x_{kj1} - \bar{x}_{kj} \\ \vdots & \ddots & \vdots \\ x_{2jT} - \bar{x}_{2j} & \cdots & x_{kjT} - \bar{x}_{kj} \end{bmatrix}$$

(9) $\tilde{\mathbf{y}}_j = \tilde{\mathbf{X}}_j^{*}\boldsymbol{\beta}^{*} + \tilde{\mathbf{u}}_j$ (das intercept α_j in Gl. 1b fällt wegen der Zentrierung weg)

Für die Schätzung $\boldsymbol{\beta}^{*}$ werden wieder die Vektoren und Matrizen gestapelt zum $NT \times 1$ Vektor $\tilde{\mathbf{y}} = \begin{bmatrix} \tilde{\mathbf{y}}_1 \\ \tilde{\mathbf{y}}_2 \\ \vdots \\ \tilde{\mathbf{y}}_N \end{bmatrix}$ (und entsprechend $\mathbf{u}_w$ [w wegen within]) und der $NT \times K$ Matrix $\tilde{\mathbf{X}}^{*} = \begin{bmatrix} \tilde{\mathbf{X}}_1^{*} \\ \tilde{\mathbf{X}}_2^{*} \\ \vdots \\ \tilde{\mathbf{X}}_j^{*} \end{bmatrix}$, so dass sich $\boldsymbol{\beta}^{*}$ wie folgt mit OLS schätzen lässt

---

[21] Zu einer Darstellung der Schätzgleichungen mit Blockmatrizen und -vektoren nach Art der Darstellung von Eckey vgl. den Anhang.

[22] Das * bei $\boldsymbol{\beta}$ und $\mathbf{X}$ soll bedeuten, dass ein um α (oder die α_j) bzw. hierauf bezogene Daten reduzierter Vektor bzw. reduzierte Matrix betrachtet wird.

(10) $\hat{\boldsymbol{\beta}}^* = \left(\tilde{\mathbf{X}}^{*'}\tilde{\mathbf{X}}^*\right)^{-1}\tilde{\mathbf{X}}^{*'}\tilde{\mathbf{y}}$.[23]

Für jedes j muss gelten (Schwerpunkteigenschaft des arithmetischen Mittels)

(11) $\bar{y}_j = \hat{\alpha}_j + \hat{\beta}_2\bar{x}_{2j} + ... + \hat{\beta}_k\bar{x}_{kj}$,

so dass sich im zweiten Schritt nach den Koeffizienten $\hat{\beta}_i$ (i = 2, …, k) gem. Gl. 10 auch die objektspezifischen α Koeffizienten schätzen lassen.

5.2. Schätzung mit dem Least squares dummy variables (LSDV) Verfahren

Im Unterschied zu Gl. 9 fallen die intercepts α_j nicht wegen Zentrierung weg. Für alle N Objekte zusammengefasst erhält man aus Gl. 1b zusammengefasst

(12) $$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} \mathbf{i} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{i} & \mathbf{i} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{i} & \mathbf{0} & \cdots & \mathbf{i} \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \Delta_2 \\ \vdots \\ \Delta_N \end{bmatrix} + \begin{bmatrix} \mathbf{X}_1^* \\ \vdots \\ \mathbf{X}_N^* \end{bmatrix} \cdot \boldsymbol{\beta}^* + \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_N \end{bmatrix} = \mathbf{D\alpha} + \mathbf{X}^*\boldsymbol{\beta}^* + \mathbf{u}\,.$$

Die Schätzung dieses Modells läuft darauf hinaus, dass man für die Unterschiedlichkeit der α Koeffizienten der N Objekte N–1 Dummies einführt, $D_2 = \begin{cases} 1 & \text{wenn } j = 2 \\ 0 & \text{sonst} \end{cases}$, …, D_N mit den Koeffizienten Δ_j für diese N–1 Regressoren. Man erhält die α Koeffizienten mit $\alpha_1 = \alpha$, $\alpha_2 = \alpha + \Delta_2$, …, $\alpha_N = \alpha + \Delta_N$ und deshalb spricht man auch bei $\hat{\boldsymbol{\beta}}^*$ vom Least squares **dummy** variables (LSDV) estimator. Zur Schätzgleichung in kompakter Matrixschreibweise vgl. Anhang.

5.3. F-Test der Hypothese ungleicher α Koeffizienten

Mit OLS erhält man bei der Schätzung des (N+K) × 1 Vektors von $\boldsymbol{\beta}^\# = \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta}^* \end{bmatrix}$ gem. Gl. 12 den Vektor $\hat{\mathbf{u}}_\alpha$ der geschätzten Störgrößen (bzw. $\hat{\mathbf{u}}_w$ bei Gl. 10), der wie $\mathbf{u}$ ein NT × 1 Spaltenvektor ist. Die F verteilte Prüfgröße (analog zur Gl. 5) zur Prüfung der Hypothese

$$H_0\text{: } \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} \neq \begin{bmatrix} \alpha \\ \vdots \\ \alpha \end{bmatrix} \text{ und } \boldsymbol{\beta}_j^* = \begin{bmatrix} \beta_{j2} \\ \vdots \\ \beta_{jk} \end{bmatrix} = \boldsymbol{\beta}^* = \begin{bmatrix} \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \quad \forall j = 1,...,N$$

also ungleicher (objektbezogener) intercepts aber gleicher Regressionskoeffizienten β_2, …, β_k ist dann gegeben mit ($\hat{\mathbf{u}}_\alpha'\hat{\mathbf{u}}_\alpha$ ist wieder die Summe der quadrierten Residuen im restringierten Modell und $S_{\hat{u}\hat{u}} = \sum_{j=1}^{N}\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j$ im nicht- restringierten Modell [vgl. Abschn. 2])

[23] Zu einer Weiterführung der Betrachtung von Gl. 9 und Interpretation von $\tilde{\mathbf{X}}^*$ als eine transformierte Matrix, die um den Mittelwert bereinigt (*demeaned*) ist, vgl. Greene, S. 561.

$$(13) \qquad F_\alpha = \frac{\left(\hat{\mathbf{u}}_\alpha{}'\hat{\mathbf{u}}_\alpha - \sum_{j=1}^{N}\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j\right) \Big/ [(N-1)k]}{\sum_{j=1}^{N}\hat{\mathbf{u}}_j'\hat{\mathbf{u}}_j \Big/ (NT - Nk)}.$$

Wegen der (im Vergleich zur Poolregression, Abschn. 3) erforderlichen zusätzlichen Schätzung von N Größen (α_1, …, α_N) anstelle von einer (nämlich das für alle j gleiche α) bei der Poolregression hat man jetzt bei F_α (im Unterschied zu F_p) im Zähler N–1 Freiheitsgrade weniger.

6. Andere Modelle und EViews

6.1. Andere Modelle

Modell mit unterschiedlichen (objektspezifischen) β aber gleichen α Koeffizienten

Es gilt jetzt H_0: $\alpha_1 = \ldots = \alpha_N = \alpha$ aber $\boldsymbol{\beta}_1^* \neq \ldots \neq \boldsymbol{\beta}_N^*$ und die Gl. 1a entsprechende Gleichung ist

$$\text{dann } \mathbf{y}_j = \begin{bmatrix} y_{j1} \\ y_{j2} \\ \vdots \\ y_{jT} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \cdot \alpha + \begin{bmatrix} x_{2j1} & \cdots & x_{kj1} \\ \vdots & \ddots & \vdots \\ x_{2jT} & \cdots & x_{kjT} \end{bmatrix} \cdot \begin{bmatrix} \beta_{j2} \\ \vdots \\ \beta_{jk} \end{bmatrix} + \begin{bmatrix} u_{j1} \\ \vdots \\ u_{jT} \end{bmatrix} = \mathbf{i}\cdot\alpha + \mathbf{X}_j^*\boldsymbol{\beta}_j^* + \mathbf{u}_j$$

und für alle N Objekte erhält man

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} \mathbf{i} & \mathbf{X}_1^* & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{i} & \mathbf{0} & \mathbf{X}_2^* & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{i} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_N^* \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \boldsymbol{\beta}_1^* \\ \vdots \\ \boldsymbol{\beta}_N^* \end{bmatrix} + \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_N \end{bmatrix} = \mathbf{X}^b\boldsymbol{\beta}^b + \mathbf{u}$$

Die OLS Schätzung liefert jetzt (in Analogie zu Gl. 2, 4 und 10)[24]

$$(13) \qquad \hat{\boldsymbol{\beta}}^b = \begin{bmatrix} \hat{\alpha} \\ \hat{\boldsymbol{\beta}}_1^* \\ \vdots \\ \hat{\boldsymbol{\beta}}_N^* \end{bmatrix} = \left(\mathbf{X}^{b\prime}\mathbf{X}^b\right)^{-1}\mathbf{X}^{b\prime}\mathbf{y}\,.$$

sowie den NT × 1 Vektor der geschätzten Residuen $\hat{\mathbf{u}}_\beta$ analog zu $\mathbf{u}$ in Gl. 1e. Mit der restringierten (bei Geltung von H_0) Residuenquadratsumme $S^0_{\hat{u}\hat{u}}$, die hier $\hat{\mathbf{u}}_\beta{}'\hat{\mathbf{u}}_\beta$ beträgt und $S_{\hat{u}\hat{u}} = \sum \hat{\mathbf{u}}'\hat{\mathbf{u}}$ (gem. Abschn. 2) erhält man die Prüfstatistik F wie folgt

---

[24] Zu den Dimensionen der Matrizen und Vektoren (K = k–1)

$\mathbf{i}$	$\mathbf{0}$	$\boldsymbol{\beta}^b$	$\mathbf{X}_j^*$	$\mathbf{X}^b$	$\mathbf{X}^{b\prime}\mathbf{X}^b$ und $\left(\mathbf{X}^{b\prime}\mathbf{X}^b\right)^{-1}$
T × 1	T × K	(NK+1) × 1	T × K	NT × (Nk+1)	(NK+1) × (NK+1)

$\mathbf{y}$, $\mathbf{u}$ und $\hat{\mathbf{u}}_\beta$ auch jeweils NT × 1 und $\mathbf{X}^{b\prime}\mathbf{y}$ ergibt einen NK+1 zeiligen Spaltenvektor wie (natürlich) auch das Produkt von $\left(\mathbf{X}^{b\prime}\mathbf{X}^b\right)^{-1}$ mit $\mathbf{X}^{b\prime}\mathbf{y}$ und damit auch $\hat{\boldsymbol{\beta}}^b$.

(14) $$F_\beta = \frac{\left(\hat{\mathbf{u}}_\beta{}'\hat{\mathbf{u}}_\beta - \sum_{j=1}^{N}\hat{\mathbf{u}}_j{}'\hat{\mathbf{u}}_j\right) \Big/ (N-1)}{\sum_{j=1}^{N}\hat{\mathbf{u}}_j{}'\hat{\mathbf{u}}_j \Big/ (NT - Nk)}.$$

Weitere Modelle

Es gibt auch Modelle, bei denen (sich auf alle Objekte gleichermaßen auswirkende) zeitliche Veränderungen (time fixed effects) modelliert werden. Auch hier kann mit Dummy Variablen ($D_s = 1$ für $t = s$ und 0 für $t \neq s$) gearbeitet werden (Modelle mit time fixed effects "control for variables that are constant across entities but evolve over time" [Stock/Watson, S. 361]).[25] Diskutiert wird auch ein kombiniertes Modell "entity and time fixed effects" $y_{jt} = \alpha_j + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + \gamma_t + u_{jt}$ oder mit $\alpha_j = \alpha + \mu_j$ (ebenda, S. 362; Greene, S. 564f).

6.2. Hinweise auf EViews

Schätzung der Modelle mit EViews

EViews erlaubt es einen Datenpool als ein eignes Objekt (das **Pool-Objekt**) zu behandeln. Es muss wie jedes Objekt mit "create" geschaffen werden und benannt werden. Man kann es z. B. auf einem eigenen Arbeitsblatt in einem workfile schaffen mit Object → new object, bei "type of object" pool auswählen und ihm dann einen Namen geben. Man wird dann nach dem Cross Section **Identifier** gefragt und es ist dann vorteilhaft, wenn man seine Daten mit einem Unterstrich und nachfolgender Kennung für die Einheit benannt hat, etwa bei Daten für die drei Städte S Stuttgart, M München und D Düsseldorf. Die Variablen (Zeitreihen) Einwohner, Mieten usw. können dann wie folgt benannt werden: ein_S, ein_D, ein_M, miet_S, miet_D, miet_M, denn dann können mit _S, _D und _M alle aufgerufen und angeordnet werden.

Unter **View** kann man verschiedene Ansichten des Pools generieren, z. B. ein **spread sheet** wenn man bei "list of … pool … series" angibt ein? miet? dann erhält man eine Tabelle mit den Spalten Einwohner und Mieten, wobei zeilenweise die Daten für D, M und S angeordnet sind. Man kann aus dieser Tabelle heraus mit der Taste **Estimate** Schätzungen vornehmen. Es kommt dann ein Menu Specification[26]/Options, worin dann anzugeben ist bei dependent variable y (etwa miet?) und bei "Regressors and AR( ) terms Common coefficients" (ein Fenster, das Eintragungen verlangt) x_1 = Miete, x_2 = … (etwa c ein?).[27]

Beim **fixed effect model** (Abschn. **) würde man die Konstante c (also das je nach Objekt unterschiedliche $\alpha = \beta_0$) nicht im Fenster "Common coefficients" mit aufführen, sondern im Fenster "Cross section specific coefficients".

Es ist dann auch nicht schwer zu sehen, wie man das Modell mit **unterschiedlichen β Koeffizienten** aber für alle Objekte gleichem α (Abschn. 4 und 5) schätzt: die Konstante c ist im Fenster "Common coefficients" und die Variablen x_1, x_2, … (also ein? usw.) sind im Fenster "Cross section specific coefficients" anzugeben.

---

[25] In Stock/Watson, S. 354 wird der Modelltyp mit dem einfachen Fall von nur T = 2 eingeführt. Man kann dann auch y und alle x Variablen in Differenzen Δy und Δx_i ausdrücken und regressieren ("before and after" regression). Im Anhang wird gezeigt, dass solche Varianten des Panel Modells in EViews möglich sind. Sie werden hier jedoch nicht weiter behandelt.

[26] Ein Menu, das nachfolgend beschrieben wird, und in dem Angaben über das zu schätzende Modell (Variablen, Fixed oder Random Effects etc.) abgefragt werden.

[27] *Alle* Variablen (außer y) im Fenster "Common coefficients" aufzuführen entspricht dem Modell der Poolregression (Abschn. 3).

EViews kann auch "**Period specific coefficients**" bestimmen im Rahmen eines Modells, bei dem sich Regressionskoeffizienten (genauer der Koeffizient α für das Intercept) von Periode zu Periode (bei allem Objekten in gleicher Weise) verändert. Es werden also Koeffizienten für die time dummies bestimmt.

Anhang

Darstellung mit Blockmatrizen

Poolregression (Kompakte Darstellung für alle Gleichungen nach Gl. 3)

(3a) $\underset{NT\times 1}{\mathbf{y}} = \underset{NT\times k}{\mathbf{X}} \cdot \underset{k\times 1}{\boldsymbol{\beta}} + \underset{NT\times 1}{\mathbf{u}}$

Die Regressoren sind in der $k \times NT$ Matrix $\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_N \end{bmatrix}$ zusammengefasst.

Distinct intercepts: Man kann Gl. 11 $\mathbf{y} = \mathbf{D}\boldsymbol{\alpha} + \mathbf{X}^*\boldsymbol{\beta}^* + \mathbf{u}$ mit der Blockmatrix[28] $\mathbf{X}^\# = [\mathbf{D} \quad \mathbf{X}^*]$ und dem $(N+k-1) \times 1$ Blockvektor $\boldsymbol{\beta}^\# = \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta}^* \end{bmatrix}$ wie folgt kompakt darstellen

$\mathbf{y} = \mathbf{X}^\#\boldsymbol{\beta}^\# + \mathbf{u} = [\mathbf{D} \quad \mathbf{X}^*] \cdot \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta}^\# \end{bmatrix} + \mathbf{u}$ so dass mit $\underset{(k-1)\times NT}{\mathbf{X}^{\#}{}'} = \begin{bmatrix} \mathbf{D}' \\ \mathbf{X}^*{}' \end{bmatrix}$ die OLS Schätzung zu

$$\hat{\boldsymbol{\beta}}^\# = \left(\mathbf{X}^\#{}'\mathbf{X}^\#\right)^{-1}\mathbf{X}^\#{}'\mathbf{y} = \begin{bmatrix} \mathbf{D}'\mathbf{D} & \mathbf{D}'\mathbf{X}^* \\ \mathbf{X}^*{}'\mathbf{D} & \mathbf{X}^*{}'\mathbf{X}^* \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{X}^*{}' \\ \mathbf{D}' \end{bmatrix} \cdot \mathbf{y}$$

führt. Wir verzichten darauf, den recht komplizierten Ausdruck für die inverse Blockmatrix anzuschreiben. Es lohnt sich aber, die Dimensionen der Matrix (Vektor) Produkte zu betrachten:

D'D ist eine $N \times N$ Diagonalmatrix mit NT in der Hauptdiagonale, also $\begin{bmatrix} NT & 0 & \cdots & 0 \\ 0 & NT & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & NT \end{bmatrix}$,

entsprechend gilt $\underset{N\times(k-1)}{\mathbf{D}'\mathbf{X}^*}$, $\underset{(k-1)\times N}{\mathbf{X}^*{}'\mathbf{D}}$ und $\underset{(k-1)\times(k-1)}{\mathbf{X}^*{}'\mathbf{X}^*}$ und somit $\underset{(N+k-1)\times(N+k-1)}{\mathbf{X}^\#{}'\mathbf{X}^\#}$ für die Momentenmatrix.

Ferner ist $\underset{(k-1)\times 1}{\mathbf{X}^*{}'\mathbf{y}}$ und $\underset{N\times 1}{\mathbf{D}'\mathbf{y}}$ so dass $\begin{bmatrix} \mathbf{X}^*{}' \\ \mathbf{D}' \end{bmatrix} \cdot \mathbf{y} = \begin{bmatrix} \mathbf{X}^*{}'\mathbf{y} \\ \mathbf{D}'\mathbf{y} \end{bmatrix}$ ein Spaltenvektor mit N+k–1 Zeilen ist, wie auch $\hat{\boldsymbol{\beta}}^\#$. Mit OLS erhält man den Vektor $\hat{\mathbf{u}}_\alpha$, der wie $\mathbf{u}$ ein $NT \times 1$ Spaltenvektor ist.

[28] Die Blockmatrix $\mathbf{X}^\#$ ist "partitioned by columns". Die Dimensionen sind bei den Matrizen (Vektoren) wie folgt (mit $K = k-1$)

D	**D'**	$\mathbf{X}^*$	$\boldsymbol{\beta}^*$	**y** und **u**	**i** und **0**	$\boldsymbol{\alpha}$
$NT \times N$	$N \times NT$	$NT \times K$	$K \times 1$	$NT\times 1$	$T \times 1$	$N\times 1$